

Methods of Evaluating the Effects of Coding on SAR Data

Melanie Dutkiewicz Ian Cumming
MacDonald Dettwiler and Associates
Richmond, B.C., Canada, V6V 2J3.

March 20, 1992

Abstract

It is recognized that mean square error (MSE) is not a sufficient criterion for determining the acceptability of an image reconstructed from data that has been compressed and decompressed using an encoding algorithm. In the case of Synthetic Aperture Radar (SAR) data, it is also deemed to be insufficient to display the reconstructed image (and perhaps error image) alongside the original and make a (subjective) judgment as to the quality of the reconstructed data. In this paper we suggest a number of additional evaluation criteria which we feel should be included as evaluation metrics in SAR data encoding experiments. These criteria have been specifically chosen to provide a means of ensuring that the important information in the SAR data is preserved. The paper also presents the results of an investigation into the effects of coding on SAR data fidelity when the coding is applied in (a) the signal data domain, and (b) the image domain. An analysis of the results highlights the shortcomings of the MSE criterion, and shows which of the suggested additional criterion have been found to be most important.

1 Motivation for the Study

Since the launch of ERS-1 in July 1991, SAR data has been received in large volumes on a routine basis (approximately 200 GBytes per day). In addition, 4 other SAR satellite sensors are planned for launch in the next decade. Data compression is needed because of the limitations of on-board tape recorders, downlinks, ground archives and data dissemination networks.

SAR data merits study on its own, because its statistical properties are quite different from those of data from other sensors. Most notably, it has a greater high frequency content and a lower signal to noise ratio (SNR) than most other sensors. Thus, compression algorithms which are optimal for other sensors may require modification before the best performance is achieved for SAR data. Also, evaluation metrics which suffice for other data sets may be insufficient for SAR data.

The main distinguishing feature of SAR data is that it is received at the satellite in the form of a hologram (the information governing the spatial distribution of energy in the image is stored in the phase of the received data). This initial received or "signal" data has a frequency spectrum similar

to that of white noise. A numerically intensive convolutional process is required to form the image, *i.e.* convert the signal data into a recognizable scene. The image formation is done on the ground.

SAR signal and SAR image data have quite different statistical properties, particularly with regard to their dynamic range, sample-to-sample correlation, and overall redundancy. For this reason and, for operational reasons, the encoding of SAR data must be studied in two distinct parts, (a) signal data encoding and (b) image data encoding. It is likely that the optimal encoding algorithm will be different in the two cases, and that the encoding performances will be quite different. Also, metrics which are used for analysing the effects of encoding on images may not be suitable for analysing the effects of encoding in the signal data domain.

2 Evaluation Metrics

The evaluation environments for SAR signal and SAR image data are given in Figure 1. As the figure shows, the analysis of SAR signal data encoding has been divided into two parts. First, the effects of encoding in the signal domain will be analysed in the signal data domain itself, by comparing original and reconstructed signal data sets using the metrics listed in Section 2.1. Second, the original and reconstructed signal data sets will each be processed, and the “original” and “reconstructed” images thus obtained will be analysed and compared in the image domain using the metrics given in Section 2.2.

2.1 Signal Domain Metrics

The properties chosen for evaluation in the signal data domain are given below. Note that each measurement is applied to the original and reconstructed data sets and the results are compared, *or* applied directly to the error data set.

1. Data Statistics

Comparison of data range, mean, standard deviation, kurtosis and entropy of the original and reconstructed data sets may highlight changes in the data characteristics. Analysis of the error data statistics (*e.g.* MSE) will show the characteristics of the error signal.

2. SQNR

The signal to quantization noise ratio is a measure of the SNR due to encoding. It will be expressed in two ways: the ratio of peak signal to MSE in dB (PSQNR); and the ratio of average signal energy to MSE in dB (ASQNR). MSE gives the total absolute encoding error between original and reconstructed data sets. ASQNR and PSQNR give relative measures useful for comparing results of encoding between data sets with differing mean and peak values respectively.

3. Data Histograms

Comparison of the histograms of the original and reconstructed data sets will show alterations in the probability distribution brought about by coding. The distribution of the error will be shown in the error data set histogram.

4. Phase Error

Phase information is used by the processing algorithm to focus the SAR image. A measurement of the phase error present in the reconstructed signal data set will allow comparison of the amount of focusing error introduced by different encoding algorithms.

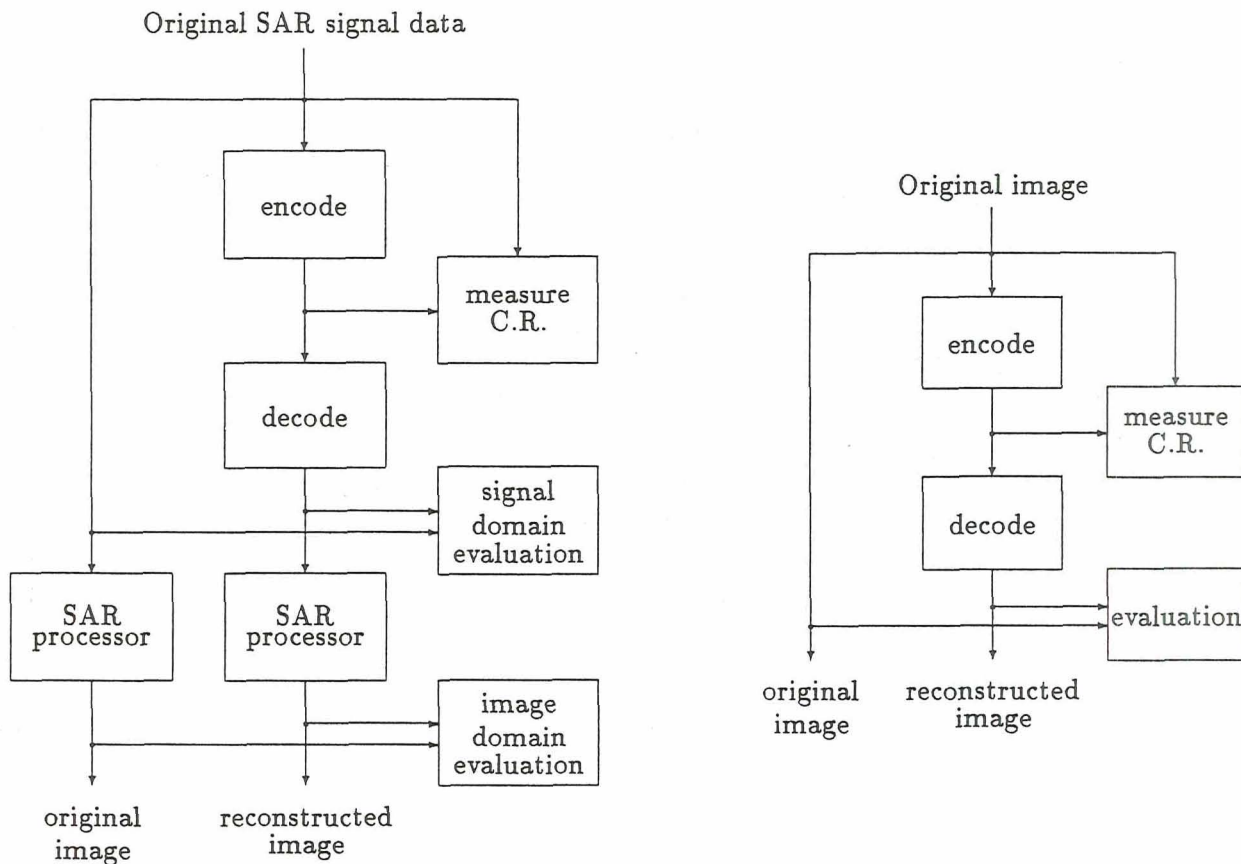


Figure 1: Evaluation environments for SAR signal data encoding (left) and SAR image data encoding (right)

Many of the evaluation metrics which are applicable in the image domain (see Section 2.2) are not applicable in this domain, *e.g.* since the appearance of signal data sets is equivalent to white noise, the effects of added quantization noise on this data will not be discernible, thus no information would be gained by examining the signal data images. These metrics will however be applied in the second part of the signal data analysis, *i.e.* after processing of the original and reconstructed signal data sets.

2.2 Image Domain Metrics

The metrics used in the image domain include items 1, 2 and 3 given in Section 2.1. Item 4 is not relevant in this case, since 4 look *detected* imagery, *i.e.* *real* data, is assumed. Phase error would however be relevant for single look complex imagery. In addition to the three given metrics, the metrics listed below will be used.

1. Image Appearance

Comparison of the reconstructed image with the original image will show how well the visual quality of the image is preserved by the coding process. The appearance of the error image will also be analysed for structural content, to see if the error is correlated with any specific feature type in the original image, and for spatial effects such as error propagation.

2. Spectra

Both the 1D and 2D spectra of the reconstructed image will be compared with that of the original image, and the 1D and 2D spectra of the error image will be examined. The latter should highlight any periodicities in the error, and will also show whether the error is greater for high or for low frequency image components.

3. Image Registration

Cross correlation between the original and reconstructed data sets, and between small corresponding regions within these data sets, will show if the coding has given rise to any mis-registration effects, either on a global or on a local scale, respectively. The latter could be important in target location and tracking, or in mapping applications, for instance.

4. Image Radiometry

The changes in the radiometric properties: mean, standard deviation, radiometric resolution and speckle statistics, brought about by coding will be measured. The radiometric linearity of the coding process will also be determined, by plotting the mean values of homogeneous regions in the reconstructed image against the mean values of the same homogeneous regions in the original images.

5. Point Target Characteristics

The position, resolution and energy distribution of point targets will be computed and compared for the original and reconstructed data sets.

6. Wavelength Estimation

The magnitude and frequency of the main peaks in the 1D range and azimuth spectra of the original and reconstructed images will be compared.

It is recognized that this list is not exhaustive, however it is felt that it is sufficient to give a detailed insight into not only the magnitude of the error, but the characteristics of the error. Figures relating to these metrics will inform the scientist using the reconstructed SAR data what information may have been altered in the encoding process and to what extent. It is important that any end user of a reconstructed SAR data should have access to such information in order to decide how much trust can be put in conclusions drawn the data.

Additional metrics which could be considered are: edge analysis; and application specific criteria such as effect on segmentation algorithms, or on ship or ice tracking algorithms.

3 Characteristics of Data Used in Study

The data used in this study was data derived from that obtained from the 1978 SEASAT SAR sensor. The major characteristics of this sensor are [Jor80]:

- radar wavelength of 0.235 m,
- incidence angles between 19deg and 25deg,
- nominal altitude of 800 km,
- noise equivalent σ_0 of -21 ± 5 dB (this implies that the SNR is in the 0 to 10 dB region),
- signal data digitized to 5 bit real samples at an IF of 11.25 MHz,
- resolution of 25 m with 4 looks.

SAR Signal Data

For the signal data encoding studies, the SEASAT data was first downconverted from real samples at the IF of 11.25 MHz to complex samples at baseband, since present day SAR sensors (e.g. ERS-1) prefer to record complex baseband data. The components of the complex data are the In-phase (I) and Quadrature (Q) components. Prior to basebanding, the data was scaled to make more use of the 8 bits available after basebanding. Thus the characteristics of the SAR signal data are as follows:

- SEASAT, downconverted to baseband,
- data precision at baseband: 8 bits per I and Q sample,
- data entropy: approx 5 bits,
- data type: signed byte,
- data range: [-128,127],
- data properties:
 - approximately zero-mean, Gaussian distribution,
 - complex data samples (I and Q components),
 - no I, Q sample correlation,
 - low inter sample correlation,
 - slow variation of *rms* level with range,
 - slow variation of *rms* level with azimuth.
- signal data size: 5100 lines \times 1250 complex samples (as required to give a 4 look detected image of 512×512 pixels).

SAR Image Data

For the image data encoding studies, the signal data was processed using a precision SAR processor. The image was processed to 4 looks. The characteristics of the SAR image data are as follows:

- SEASAT, 4-look detected, resolution 25m,
- data precision: 12 bits per pixel,
- data entropy: approx 10 bits,
- data type: unsigned word,
- data range: [0,4095],
- data properties:
 - Rayleigh amplitude distribution
 - very high dynamic range:
 - * >50 dB for point targets,
 - * >30 dB for distributed areas,
 - statistics dominated by speckle:
 - * moderate inter-sample correlation,
 - * less spectral roll-off than optical sensors,
 - * lower SNR than optical data,
- single band,
- image size: 512×512 pixels.

4 Selection of Algorithms

Two encoding algorithms have been used in each part of the investigation:

1. for encoding in the signal domain, JPL's Block Adaptive Quantization (BAQ) algorithm (as used on the Magellan mission to Venus), and Unstructured (full search) Vector Quantization (UVQ) have been compared,
2. for encoding in the image domain, the JPEG DCT algorithm and UVQ have been compared.

The BAQ [KJ89] algorithm was selected for analysis because of its simplicity and because it has given very good results for SAR images obtained on the Magellan mission. The algorithm has been designed specifically for the zero-mean Gaussian statistics of SAR signal data, and is therefore not suitable for the Rayleigh distribution of SAR image data. In its present form, BAQ provides a fixed compression ratio of 4 (*i.e.* from 8 bpp to 2 bpp).

The VQ [Gra84] algorithm was selected because it is the subject of much of the current research in encoding, and it has been shown to give excellent performance on some data sets. The decision to use unstructured VQ was reached because it was believed to give better performance than structured versions, which generally compromise quality for a reduction in "cost" of execution. In retrospect, it would have been better to use a structured version, such as gain/shape VQ, since the cost of execution of UVQ was found to be so high that its full potential was unrealizable, and compromises had to be made, such as working with smaller vectors.

The UVQ algorithm was applied in both the image and signal data domains.

The JPEG DCT algorithm [Wal90] was selected as the second algorithm to be applied in the image domain, since it is becoming a globally accepted still image compression standard. It is recognized that this algorithm is not optimized for SAR data, for example, compression results can be improved using an adaptive DCT algorithm such as that of Chen and Smith [CS77]. The JPEG DCT algorithm was not applied to SAR signal data since the JPEG quantization tables as they stand are not applicable to complex data. SAR signal data has a spectrum similar to that of white noise, with almost as much energy in the high frequency components as there is in the low frequency components, and it is unlikely that the DCT algorithm would produce sufficient energy compaction to enable a useful level of data compression without appreciable loss of image fidelity.

5 Methodology of Study

The activities of the study are broken down into the following steps:

1. Select a suitable operating point *i.e.* compression ratio, and
2. do a full comparison of the algorithms, using all the defined metrics, at the chosen operating point.

To select the operating point for the image data encoding analysis, the UVQ and JPEG DCT algorithms were run with a variety of compression ratios, and SQNR was graphed against compression ratio. The images obtained at each compression ratio setting were viewed, and an operating compression ratio of ≈ 7 was selected, since in this region there was felt to be a small but acceptable

level of distortion in the data. The operating parameters for UVQ were 2×2 vectors and 2^7 codevectors. For JPEG DCT, the quality parameter was set to 8.

For the signal domain encoding, UVQ was run at a variety of compression ratios, and the resulting SQNR compared with that of BAQ at its fixed compression ratio of 4. Since UVQ algorithm produced a similar SQNR to BAQ at the same compression ratio, the operating compression ratio for the signal data study was set at 4. The operating parameters for UVQ for this compression ratio are: 2×2 vectors and 2^8 codevectors.

Three SAR scenes, which provided a variety of scene features, were used for the analysis:

1. Cambridge Bay (sea ice, ocean, snow and arctic tundra),
2. Niagara Falls (urban, roads, river, agriculture, airfield, orchards), and
3. Flevoland (sea, polder, forests, urban).

6 Signal Data Encoding Results

The results of the encoding experiment on the SAR signal data pertaining to the Niagara image, using the BAQ and UVQ algorithms, are presented in Table 1. The Niagara results have been selected for discussion since they accentuate the differences in the results for the two algorithms. The other scenes produced similar results, but the differences were less marked. The reason for this being that the Niagara scene contained greater feature diversity than the other two scenes.

The results for the analysis in the signal (and coding) domain show that BAQ produces a lower SQNR, and in general alters the signal data statistics more than the UVQ algorithm. BAQ does however produce a lower average phase error than UVQ.

Both algorithms altered the appearance of the signal data histograms. For BAQ, the reconstructed histogram shows 4 main peaks, corresponding to the data being mapped to $-GB_2$, $-GB_1$, $+GB_1$, or $+GB_2$ where G is a function of the *rms* level of a block, and the B_i values are chosen to minimize the *rms* error ($B_1 \approx 0.5$, and $B_2 \approx 1.7$). The concentration of the data in 4 main peaks implies that the *rms* level varies very little from block to block. For UVQ, the histograms, which were smooth Gaussian distributions on input, appeared very spiky on output, but still retained the same overall Gaussian appearance. The spikiness is due to the fact that the use of a limited size codebook restricts the output intensity space, e.g. by limiting the combinations of intensities of adjacent pixels.

The results for the analysis in the image domain (*i.e.* after processing both original and reconstructed signal data sets) show that the SQNR is again lower for BAQ than for UVQ. In this domain, however, BAQ has degraded the data statistics less than UVQ.

The original, reconstructed and error images for BAQ and UVQ are shown in Figure 2. The reconstructed images for BAQ and UVQ are judged to be of good and equivalent quality. The error images show the error to be fairly evenly distributed for BAQ. For UVQ, the structure of the original image is fairly evident in the error image. This implies that the algorithm is affecting some features more than others. It is seen that the error image is in essence a negative of the original, implying that bright areas have become darker, and dark areas brighter. For BAQ this occurs to a much lesser extent. The phenomenon implies that the radiometric linearity of the coding has been disturbed. This last deduction is corroborated by the fact that the linearity is measured to be 0.98 for BAQ and 0.94 for UVQ, *i.e.* despite the higher MSE value, BAQ has disturbed the linearity

less than UVQ. Preservation of linearity is important, for example, if the data is to be used for SAR sensor calibration.

In terms of the other image domain metrics, both algorithms performed very well, *i.e.* image histograms, spectra, radiometric resolution, speckle and point target characteristics were well preserved (*i.e.* the error is too small to be statistically significant). BAQ did however perform better for wavelength estimation, preserving both the magnitude and frequency of the dominant peaks in the 1D spectra. UVQ preserved the frequency but reduced the magnitude by $\approx 6\%$.

To summarise, when evaluated in both the signal and image domains, BAQ gave a higher MSE than UVQ. However, since BAQ gave better results in terms of data statistics, radiometric linearity and wavelength estimation, and equivalent results for the remaining metrics, it is deemed to be better than UVQ, with the given operating parameters, for application to SAR signal data encoding.

It is recognized that gain/shape VQ would result in improved linearity, but for the same compression ratio and vector size, this would be at the expense of codebook size. Thus one would expect the overall MSE to be essentially the same, but with the errors more evenly distributed. This last point illustrates another of the shortcomings of MSE: since it is a global measure, it gives no indication of the spatial distribution of the error, and no indication of the extent to which the error affects the various properties important to SAR image analysis.

7 Image Data Encoding Results

The results of encoding experiments on the SAR image data for the Niagara scene, using the JPEG DCT and UVQ algorithms, are presented in Table 2. Again, the Niagara results have been selected for discussion since they accentuate the differences in the results for the two algorithms.

The results for the analysis in the image (and coding) domain show that the SQNR is lower for UVQ than for JPEG DCT, and, in general, UVQ has distorted the data statistics more than the JPEG DCT.

The original, reconstructed and error images for JPEG DCT and UVQ are shown in Figure 3. The reconstructed images for both JPEG DCT and UVQ are judged to be of good quality, though the JPEG DCT reconstructed image is slightly less well focussed than the original (due to attenuation of the high frequency components), and the UVQ reconstructed image shows some loss of contrast (due to the codebook having a more limited intensity space than the original data). The error images show the error to be fairly evenly distributed for both algorithms, with the UVQ error image having slightly more structure than the JPEG DCT error image.

The histogram characteristics were well preserved by the JPEG DCT algorithm. For the UVQ algorithm, although the overall shape of the histogram was preserved, the histogram was much more "spiky" in appearance. Again this is due to the codebook limiting both the intensity space, and the intensity combinations of adjacent pixels.

For both algorithms, the reconstructed spectra appeared very similar to the original spectra, with all dominant peaks reproduced, and no extra peaks introduced. The error spectra did not show evidence of any periodic error effects having been introduced. The error spectra did show that the error for JPEG DCT was greater in the high frequency portion of the spectrum than in the low frequency portion. This was also true for UVQ, but to a very much lesser extent.

Both algorithms preserved the radiometric resolution, and both reduced the speckle component by

$\approx 5\%$. The radiometric linearity was 1.0 for JPEG DCT, however UVQ was again shown to disturb the linearity, with the slope of the linearity plot being 0.96.

Neither algorithm produced any mis-registration effects, either globally or locally. For almost all the point targets studied, JPEG DCT was found to produce only very slight differences in peak widths, energy and position. UVQ, however, was found to be more likely to alter these characteristics. This is mainly due to the fact that point targets are essentially "rare events" of dimension not much greater than the vectors themselves. A centroid based codebook is not geared to meet the needs of reproducing the individual characteristics of each point target. Point target preservation is important in SAR image analysis.

In terms of wavelength estimation, JPEG DCT preserved both the magnitude and frequency of the dominant peaks in the 1D range and azimuth spectra. UVQ preserved the frequency, but reduced the magnitude by up to 3%. It is however recognized that the peaks studied occurred in the low frequency portion of the spectrum. For higher frequency peaks, JPEG DCT would be more likely to alter the magnitude, since the evaluation of the spectra showed that JPEG attenuates high frequency components more than low frequency components (an inherent property of the JPEG quantization scheme).

To summarise, the JPEG DCT algorithm performed consistently better than the UVQ algorithm. The use of the additional metrics specified in this paper served to highlight the exact way in which the higher error introduced by the UVQ algorithm manifested itself.

Again, the linearity of VQ could be improved using gain/shape VQ. It may also be possible to reproduce the point target characteristics more accurately using some form of classified VQ, as opposed to designing a purely centroid based codebook.

8 Conclusions

In this paper we have suggested a number of evaluation criteria especially suited to determining the manner and extent to which an encoding algorithm distorts SAR data. Both signal and image domain encoding were studied, with two algorithms having been applied in each domain. The results obtained using the given algorithms were described and analysed using the given metrics.

The measurement of MSE or SQNR is important, as it gives a measure of the severity of the unstructured error brought about by encoding. The MSE should always be included as an evaluation metric. However, in order to gain additional insight into the characteristics of the error, and also into the manner of operation of the encoding algorithm, additional metrics are needed.

The metrics which produced the most insight into the manner in which the algorithms acted on the data and the resulting degradations were felt to be:

1. the original versus reconstructed data histograms,
2. error images,
3. error spectra,
4. radiometric linearity, and
5. point target analysis.

The data histograms were useful in both the signal and image domain analysis in showing how the algorithms acted on the data. *e.g.* for BAQ they showed how little the block rms value varied

for the chosen block size. This insight would be useful when selecting a block size. For UVQ the histograms showed the effects of mapping data onto a smaller intensity space and limiting the allowed intensity combinations of adjacent pixels.

The error images, in conjunction with other metrics, such as the data histograms, showed which properties of the original images were giving rise to the highest errors *e.g.* for UVQ it was found that regions with an intensity having a low probability of occurrence, in this case those on the edges of the input intensity space - very bright areas or very dark areas - produced the largest errors.

The spectra highlighted which frequency components were most affected by the coding.

The measurement of radiometric linearity consolidated the results seen in the error image, and showed the severity of this error.

The point target analysis showed how well rare events are reproduced by the given algorithm.

9 Acknowledgments

This work is being performed under ESA Contract No. 9122/90/NL/PR(SC). The authors would like to thank the technical authority, Jean-Luc Marchand of ESTEC, for his advice and direction during the course of this study.

References

- [CS77] W.H. Chen and H. Smith. Adaptive Coding of Monochrome and Color Images. *IEEE Trans. on Communications*, COM-25(11):1285-1292, Nov 1977.
- [Gra84] Robert M. Gray. Vector Quantization. *IEEE ASSP Magazine*, 1:4-29, Apr 1984.
- [Jor80] Rolando L. Jordan. The SEASAT-A Synthetic Aperture Radar System. *IEEE Journal of Oceanic Engineering*, OE 5(2):154-164, April 1980.
- [KJ89] R. Kwok and W.T.K. Johnson. Block Adaptive Quantisation of Magellan SAR Data. *IEEE Trans. on Geoscience and Remote Sensing*, 27(4):375-383, July 1989.
- [Wal90] Gregory K. Wallace. Overview of the JPEG Still Image Compression Standard. *SPIE Image Processing Algorithms and Techniques*, 1244:220-233, 1990.

Table 1: Evaluation of Signal Encoding Performance (Niagara Scene)

Signal Domain Results	BAQ	UVQ
CR at operating point	4	4
MSE	107	76
ASQNR (dB)	9.5	11.0
PSQNR (dB)	21.8	23.3
Data range	-20 %	-27 %
Δmean	+24 %	+2 %
$\Delta\sigma$	+8 %	-4 %
$\Delta\text{kurtosis}$	-8 %	-8 %
$\Delta\text{entropy}$	-32 %	+7 %
Phase Error (degrees)	18.1	23.2
histograms	4 peaks	spiky
Image Domain Results	BAQ	UVQ
MSE	9992	7332
ASQNR (dB)	18.6	20.0
PSQNR (dB)	32.2	33.6
Δmean	+8 %	-3 %
$\Delta\sigma$	-1 %	-7 %
$\Delta\text{kurtosis}$	+0.6 %	+18 %
$\Delta\text{entropy}$	-0.1 %	-1 %
histograms	preserved	preserved
radiometric resolution	preserved	preserved
speckle	preserved	preserved
radiometric linearity	0.98	0.94
mis-registration	none	none
point targets	preserved	preserved
wavelength estimation	preserved	degraded

Table 2: Evaluation of Image Encoding Performance (Niagara Scene)

Image Domain	JPEG DCT	UVQ
CR at operating point	5.7	6.9
MSE	6208	13430
ASQNR (dB)	21.5	18.1
PSQNR (dB)	34.3	31.0
Δ range	+0.8 %	-6 %
Δ mean	-0.04 %	-0.4%
$\Delta\sigma$	-0.1 %	-3%
Δ kurtosis	-2 %	-5%
Δ entropy	+0.1 %	-19%
histograms	preserved	spiky
radiometric resolution	preserved	preserved
speckle	-5%	-5%
radiometric linearity	1.00	0.96
mis-registration	none	none
point targets	preserved	degraded
wavelength estimation	preserved	degraded

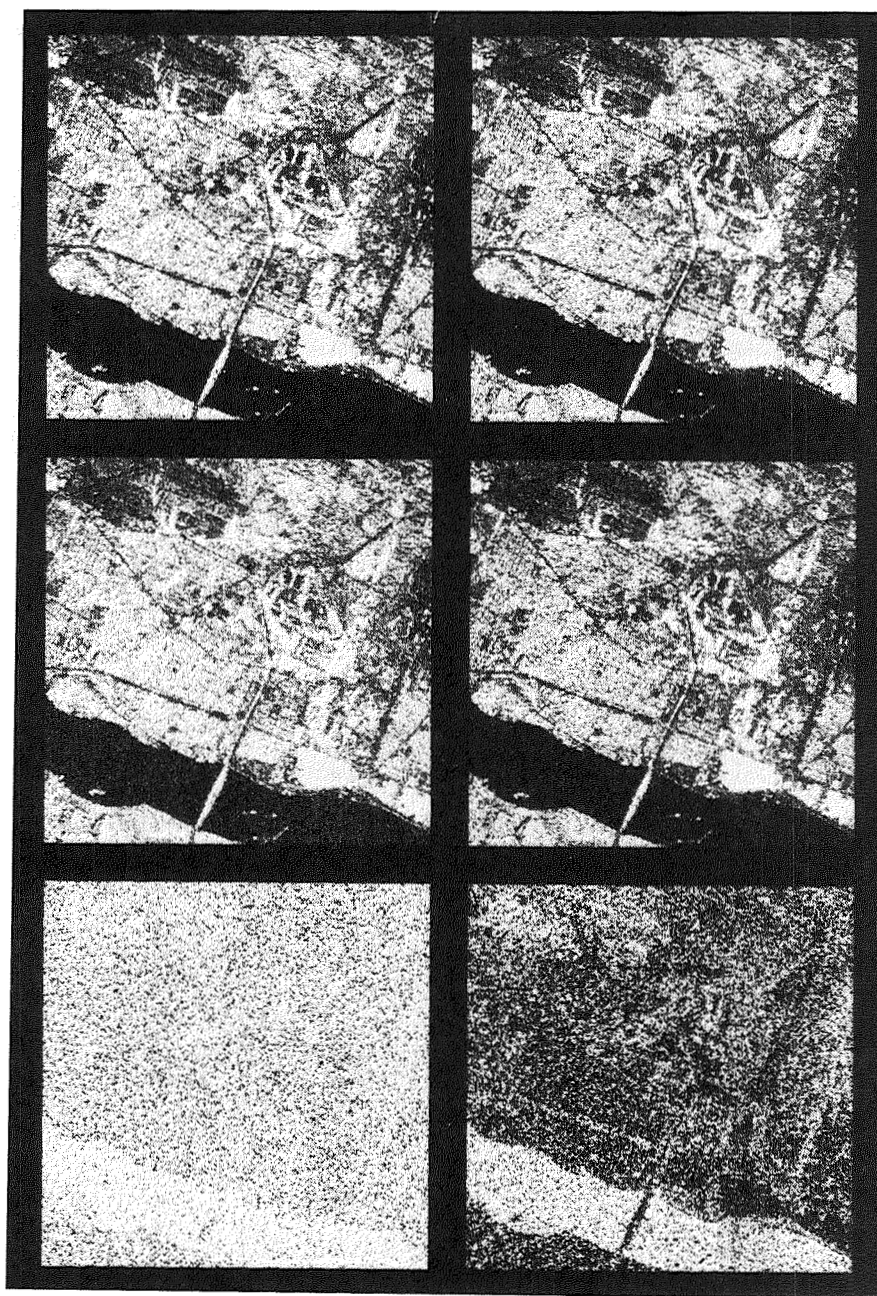


Figure 2: Niagara: Original, Reconstructed and Error Images (Signal Data Encoding)

Top Left: Original;	Top Right: Original
Centre Left: BAQ Reconstructed;	Centre Right: VQ Reconstructed
Bottom Left: BAQ Error;	Bottom Right: VQ Error

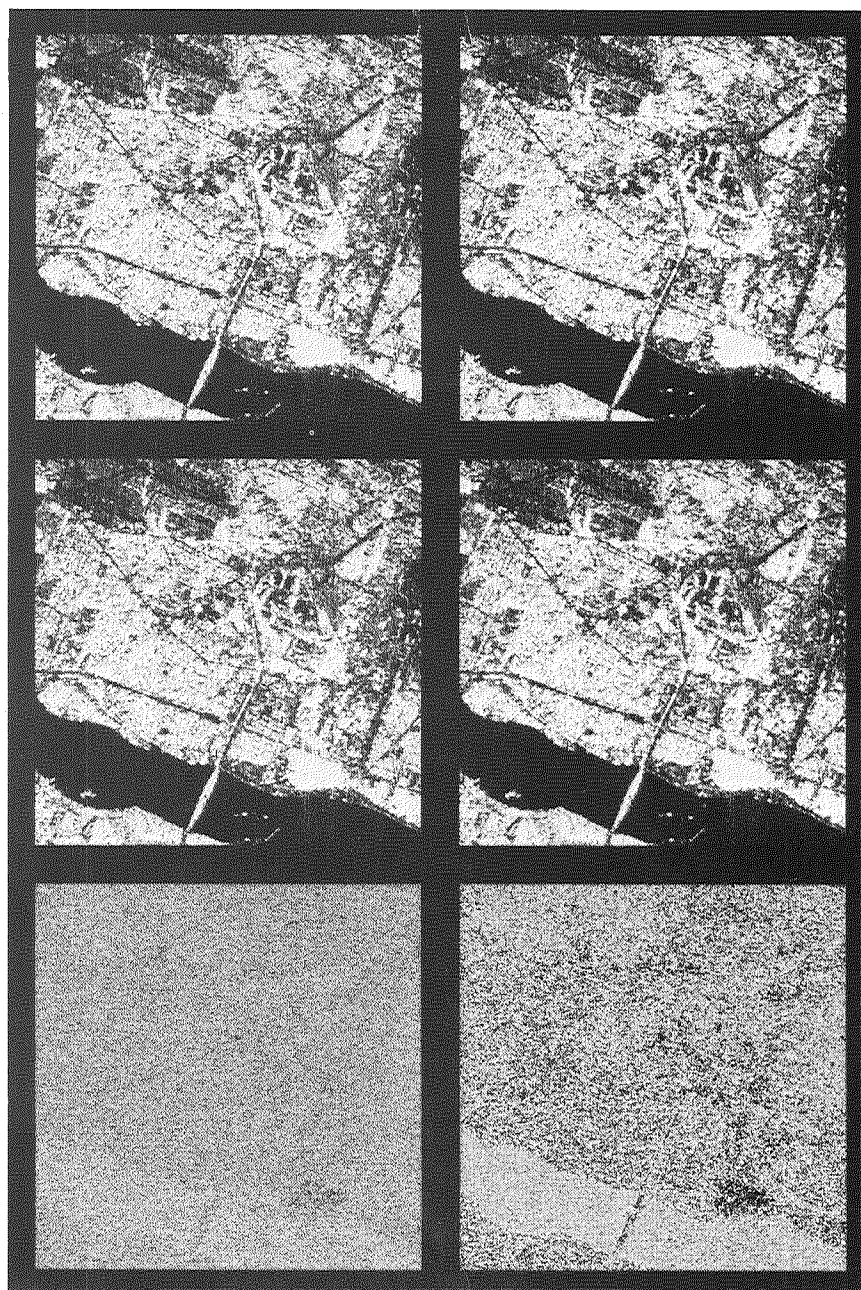


Figure 3: Niagara: Original, Reconstructed and Error Images (Image Data Encoding)

Top Left: Original;	Top Right: Original
Centre Left: JPEG DCT Reconstructed;	Centre Right: VQ Reconstructed
Bottom Left: JPEG DCT Error;	Bottom Right: VQ Error